

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 19:54:31 ; Search time 57.092 Seconds  
(without alignments)  
13439.802 Million cell updates/sec

Title: US-10-027-000-3

Perfect score: 2502  
Sequence: 1 atggtcatattgatgttga.....attgtgtctggtgtgttaa 2502

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapept 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata1/lna/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata1/lna/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata1/lna/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata1/lna/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata1/lna/PCRTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata1/lna/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	184.4	7.4	3241	US-09-434-288-11	Sequence 11, Appl
2	179	7.2	2401	US-09-320-878-20	Sequence 20, Appl
3	175.2	7.0	2430	US-09-105-537-23	Sequence 23, Appl
4	175.2	7.0	13613	US-09-105-537-3	Sequence 3, Appl
5	138.6	5.5	50937	US-09-428-517-1	Sequence 1, Appl
6	138.6	5.4	2166	US-09-134-078-5	Sequence 5, Appl
7	111.6	4.5	4403765	US-09-103-840A-2	Sequence 2, Appl
8	111.6	4.5	4411529	US-09-103-840A-1	Sequence 1, Appl
9	101.4	4.1	16836	US-09-147-236-10	Sequence 10, Appl
10	101.4	4.1	16836	US-09-147-236-10	Sequence 10, Appl
11	84.6	3.4	12588	US-08-387-942C-1	Sequence 7, Appl
12	84.6	3.4	390	US-09-197-649-7	Sequence 1, Appl
13	83.8	3.3	3033	US-08-462-080B-1	Sequence 1, Appl
14	83.8	3.3	3033	US-08-462-090-1	Sequence 1, Appl
15	83.8	3.3	3033	US-08-463-461-1	Sequence 1, Appl
16	80.4	3.2	1140	US-09-023-173-4	Sequence 4, Appl
17	78.6	3.1	2064	US-08-343-428-1	Sequence 1, Appl
18	78.2	3.1	1294	US-09-025-651-2	Sequence 2, Appl
19	78.2	3.1	1521	US-08-083-948-14	Sequence 14, Appl
20	78.2	3.1	1521	US-08-393-785-14	Sequence 14, Appl
21	78.2	3.1	1521	US-08-475-694-14	Sequence 14, Appl
22	78.2	3.1	1521	US-08-712-057-14	Sequence 14, Appl
23	78.2	3.1	1647	US-08-083-948-13	Sequence 13, Appl
24	78.2	3.1	1647	US-08-393-785-13	Sequence 13, Appl
25	78.2	3.1	1647	US-08-475-694-13	Sequence 13, Appl
26	78.2	3.1	1647	US-08-712-057-13	Sequence 13, Appl
27	78.2	3.1	1865	US-08-083-948-7	Sequence 7, Appl

28	78.2	3.1	1865	US-08-393-785-7	Sequence 7, Appl
29	78.2	3.1	1865	US-08-475-694-7	Sequence 7, Appl
30	78.2	3.1	1865	US-08-712-057-7	Sequence 7, Appl
31	78.2	3.1	2182	US-09-347-878-35	Sequence 35, Appl
32	76.2	3.0	4403765	US-09-103-840A-2	Sequence 2, Appl
33	76.2	3.0	2712	US-09-025-691-4	Sequence 4, Appl
34	74.4	3.0	1143	US-08-387-942C-19	Sequence 19, Appl
35	71.8	2.9	1320	US-08-461-775-8	Sequence 8, Appl
36	71.8	2.9	1320	US-09-031-606-8	Sequence 8, Appl
37	71.8	2.9	1620	US-08-461-775-10	Sequence 10, Appl
38	71.8	2.9	1620	US-09-031-606-10	Sequence 10, Appl
39	70.8	2.8	1929	US-09-380-420C-1	Sequence 1, Appl
40	70.8	2.8	13842	US-09-105-537-30	Sequence 30, Appl
41	70.8	2.8	36778	US-09-105-537-5	Sequence 5, Appl
42	69.6	2.8	38506	US-09-320-878-19	Sequence 19, Appl
43	69.6	2.8	2219	US-08-510-646B-17	Sequence 17, Appl
44	69.2	2.8	2668	US-08-461-775-11	Sequence 11, Appl
45	69.2	2.8	2668	US-08-031-606-11	Sequence 11, Appl

#### ALIGNMENTS

RESULT 1  
US-09-434-288-11  
; Sequence 11, Application US/09434288  
; Patent No. 6303767  
; GENERAL INFORMATION:  
; APPLICANT: Bellach C. Melanie  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA  
; FILE REFERENCE: 30062-20030.00  
; CURRENT APPLICATION NUMBER: US/09/434,288  
; PRIOR FILING DATE: 1999-11-05  
; PRIORITY APPLICATION NUMBER: 60/107,093  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 3241  
; TYPE: DNA  
; ORGANISM: Streptomyces narbonensis  
US-09-434-288-11

Query Match 7.4% Score 184.4; DB 4; Length 3241;  
Best Local Similarity 48.2%; Pred. No. 6.3e-31;  
Matches 1028; Conservative 0; Mismatches 956; Indels 150; Gaps 12;

QY	206	CGCTGGTTCACATTCACCAACTGCTCGAAGAGCGAGTAAAGATGCGCAAG	265
DB	1151	CCCTGGCCGACACCTTCGACACTCCATGCGCAGACTACAGAGGTCATGCGCGCG	1210
QY	266	AGGCATCGCTAGAGTGGCGCATGTGCTCGGCGCCGATATCAACATCAACGCTCC	325
DB	1211	ACGACCGCGGCGTGGCGACAGATGCTTGGCCCGATGATGACACATTCGGGTGC	1270
QY	326	CTCTGGTGGACGCGTTCGAGTGCATGTCGAGATTCCTCGCGGCTTGGAG	385
DB	1271	CACACGGCGCGGACATTCAGACCTTCAGCGAGAGCCCGCTGCTCTCGGACCG	1330
QY	386	CTGGCGCTTCATCCGCGCATTCAGACCTGAGTGCAGGCTAGCATTAACGACTTT	445
DB	1331	CGTGGCCGACATCAAGGCGATCCAGGCTCGGCTGATGACCCAGCGCAACGACTTC	1390
QY	446	TGTCATGATGATGAGGAGGAGCGGCGATGATGTCAGAGCATGTCAGGAGCGGCTC	505
DB	1391	CGGCAACACACAGAGAGACACCGCTTCACCGTCAACGCGCAGTGCAGAGAGAGCG	1450
QY	506	TCCGTGAATTCAGCACTCCGCTTCAGATGTGTCGAGAGTCCACCGGGTCCGT	565
DB	1451	TCCGCGAGATGAGTTCGCCGCGTTCGAG--GCGTCTCGAAGGCGCGGCGGCTCTCT	1507

QY	566	TCATGTGCGGCGATCATGGCATGATGGCCGCTGTGCACAGCAAGCCCTAAATATCTTG	625
Db	1508	TGATGTGTCCCTATTAACGGCGTCAACGGCAAGCCGCTCTCGGCAACAGACAGCTCTCA	1567
QY	626	ATGGGATGCTTGAAAGGAATGGGGTGTGGAGTGGCTTAATCATGAGCAGCTGGTACGGCA	685
Db	1568	ACAACTGCTGCGGCAACGACGATGGGGCTTCCAGGGCTGGGTGATGTCCGACTGGCTGCGCA	1627
QY	686	CATACAGTACACACAAACACCCGTTGTGGCAGCCCTTCGACTTCGAGATGCCCCGACCTCAC	745
Db	1628	C---CCCGGGCGACGGACGCTCATCAACAAAGGCGCTCGACACGGAGAT---GGGCGTCGAG	1680
QY	746	GCTTCCGAGGAAACATCACTCAAGTTTCAACGCTGCTCAACGGGAAAGGCCCTTTATTCAGCTCA	805
Db	1681	CTTCCCGGGGACATCCCGCCGGCGAGACCTTCGCGCCGCCCAAGTTTCTTCGTTGACGGG	1740
QY	806	TTGACCAGAGGGGCTTAGGAAGTTCTTGAGTTTGCTCAAGAAGTGTGCTGCTCCGGAGTGA	865
Db	1741	CTGTA-----AGCAGGCGCTCGTAAGCGCACAGCTGTCCCGAGGGG	1780
QY	866	CGGAGACGGCCCCCGACAGCAGCTGTCAACACACCCCGAAAGCGCAGCTTCTTCGCGGA	925
Db	1781	CGGTGACGGGGTGGCGGAGCGCATCTGTCAACAGATGAGACAAAGTTGCGGTGCTCTCTCG	1840
QY	926	AGGTTGGCAACGAGGGCATGTGCTGTGTAAGAAACGAAACAAAGTTTGTGCCCTTGACGA	985
Db	1841	CGACTTCGGCGCCCCCGCCGAGGCTTACAAGGCGGGCGCCACAGCGGCTGTCCGCAAGG	1900
QY	986	AGAAAGAGAGACGCTGATTTCTCGGCCCAACGCCAAGCAGGCCACATATACAGCGCGAG	1045
Db	1901	TCGCGGAGAACGGCGGGGTGCTTCTGTGGCAACAGGGGCCAGGCGCTTCGCGCGGGTG	1960
QY	1046	GCTTGTGCGGCACTAGGGCCTTACTAGCAGTCACTCCCTTTGAGCGCCTCAGACAGCAGC	1105
Db	1961	ACGCGCGGAAGAGATGCGCTCATCGGCCCGACGGCGCTCGAC--CCAAAGGTACACGGC	2019
QY	1106	TCGAGACGCGCGCAATCGTATACACGCTGGGGCGCTTACACACACCGTTCTCTCCATTCTAGGCG	1165
Db	2020	CTGGGCAAGGCCACAGCTGTCCCGGAGTCTGGCGGGCGGCGCCCTCTGTGACACCATCAAGGCC	2079
QY	1166	AGCAGTGGCTTCAACGCCGAGCGGCGCTTCCGGGCAATGGGCTGAGGGGCTTTCAACGAGCCCC	1225
Db	2080	CGCGGGGCGGGGTGGAGAGGTGACTACGAAACGGGTGAGAGAACTTTCGGGACGCGGG	2139
QY	1226	CTGGTACCCCTTAACCGCCAGACATTGACGAGCTTCTTTCACCAAGACGGAATGCACC	1285
Db	2140	ATCCCGGGGCGGCGAGCTCAGC-----CCGGGCTTACACC	2173
QY	1286	TGGTGGACTTACTACCACCCCAAGGCGGACAGACGCTGTAAGCCGACATGTAAGGGCAGCT	1345
Db	2174	AGGGGCACTACTTGAGACCGGGCAAGGCGGGGGCGCTGTATACAGCGGCAAGCTGTACCGTGC	2233
QY	1346	ACACCGGCGAGAGACTTGCACCTTACGAGCTCGGCGCTGCTCTTGCAGCGACGCAAGG	1405
Db	2234	CGCGGAGCGGAGTACTCGGATCTCGGTCAAGGCCACCGGTGGCTATCGCAGAGGTGCAG--	2292
QY	1406	CGTACGTAGACGACCACTGTGTGTGACACGCCACCAAGCAGGTCCCGGCGATGCTT	1465
Db	2293	-----CTCGGCGACCA	2305
QY	1466	TCTTCGGGCTCGGCAACCGGAGAGAAAGGCGCGATTAATTCGTCTCAAGGCGCAACAGCT	1525
Db	2306	CCATGAGGGCGGTACAGTTTACGCGCAAGGTAGACAGCCCGCTCTCTCAAGCTGTACCAAGG	2365
QY	1526	ACAAGTTCAAGATTCGAGTTTGGCTCCGCAACCCACTTACACCTTCAAGGGCGACACCAATCG	1585
Db	2366	GCAAGCAACAAGCT-----CACGATCTCGGGCTTCCCGATGA	2401
QY	1586	TCGCCGGGCAACGGCTCCCTTCGCGGCTGGCGGCTGCAGAGTATTTGACGACACCGGCCMAA	1645
Db	2402	GCGCGACGCGGCTCTCCCTTGAAGCTGGGCTGGGTGACCGCGGAGGCGACGCGACGCGACGA	2461
QY	1646	TCGAAAAGTCCGCTCGCCCTTCGCCAAGGACAGACAGCCAGTATCATCTGCGGGGCGCTTA	1705

Db	2462	TCGCGAAGGCGGTGGAGTGCAGCGGAAGGCCGCTACGGCCATCGTTCGGG-----	2514
Qy	1706	ACGCGGACTGGGAGACCGAGGGGCGCCACCGCGAGCATGAAGCTCCCCGGCGTGTGG	1765
Db	2515	--TAGAGAGAGCGACCGAGGGGCGCTGCACCGTCCGAACCTGTGGCTGCCGGGTACGACAG	2572
Qy	1766	ACCAGCTAATTGGCGAGCGTGGCGCGCGGAAACCCAAACACCGTGTGTATGACAGACGG	1825
Db	2573	ACAACTGATATCGCGCGGTGCCACAGCCGAACCCGAACAGATCGTGTCTCAACACCG	2633
Qy	1826	GCACCCCGGAGAGATGCCCTGGGTGCAGCCCAACGCCCGCGTCAATCCAGGCGCTGTACG	1882
Db	2633	GTTCGTGCGGTGTGTGTGCGGTGTGTGTCGAAGACCCGCGGGTCTGGAGCATGTGGTACC	2692
Qy	1886	GGCGGACGAGACGAGCGGCAACTCATTCCTCCGACATCGTCTTTGGCGCATCAACCCCTGG	1945
Db	2693	GGGGGACGAGCGGGGCGCGGAGGCGACCGCGCGCTGTCTACGTTAGCTGAACCCGAGCG	2755
Qy	1946	GCAAGCTCTCCCTACGCTTC-----CGAAGCGCGTGCAGAGCAACCCCGGTTTCGA	1999
Db	2753	GCAAGCTCACGACAGAGTTCGCCGCGCGCGAGAAACGACGCGCGTGGCCGCGACCCGA	2812
Qy	2000	ACTTCGCGACCGAGCGCGGGCGGACGCT---GTACGGCGAGACGCTCTACGTGCGGTACA	2056
Db	2813	ACCGGTACCGGGGCGTGACCAACCGAGAGACGTACAGCGAGGGGATTCACAGTCGGGTACC	2872
Qy	2057	GGTACTACGATTTTGCCGACAAAGAGAGCTCAATTCCCTTTGGCCACGGCGTGTCTACA	2116
Db	2873	GCTGGTTGCAACAGAGACAGCTACAGCCGCTGTTCCTCCGTCGGGACAGCGGCTGTGTACA	2922
Qy	2117	CCACTTTTGCCTTTCCATATCTCTCGGTCTCTCACAAAGAC--GACAAAGCTGACCGTGT	2173
Db	2933	CCTGCTTACGACAGAGGCGCCCGACCGGTGTGTGCGACGTCAACGGGCGGCGTGAAGGTCA	2992
Qy	2174	CCCTCTCCGTAAATAACACCGGCGTCCGTGCGCGGCGACAGGTGGCCCAAGCTTACGTCA	2233
Db	2993	CGGTACGCGTCCGCGMACAGGGGCGACGCGCGGGCGAGAGGTGTCTCCAGGCGTATCTCG	3052
Qy	2234	AGCCCTCTCAAGGCGGCCCAATTAACCGCCCGCTCAAGGAGGTCAAGGGGCTTCGCAAGG	2293
Db	3053	GGCGGAGCGCGAAGGTACGGCTCCGACGAGCGAGAAAGAAAGACTTCGTGGGCTACAGAAAG	3112
Qy	2294	TCGAAGTGCAGCCCGCGGAGACGAGCGGTGAC	2327
Db	3113	TCGGCGTCCGGCGGCGGAGTGCAGAGCGGTGAC	3146
RESULT 2			
-US-09-320-878-20			
-Sequence 20, Application US/09320878A			
-Patent No. 6117659			
-GENERAL INFORMATION:			
APPLICANT: ASHLEY, Gary			
APPLICANT: BETLACH, Melanie C.			
APPLICANT: BETLACH, Mary C.			
APPLICANT: MCDANIEL, Robert			
APPLICANT: TANG, Li			
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLKETIDE SYNTHASE			
FILE REFERENCE: 300622002120			
CURRENT APPLICATION NUMBER: US/09/320, 878A			
CURRENT FILING DATE: 1999-05-27			
EARLIER APPLICATION NUMBER: CIP OF 09/141, 908			
EARLIER FILING DATE: 1998-08-28			
EARLIER APPLICATION NUMBER: CIP OF 09/073, 538			
EARLIER FILING DATE: 1998-05-06			
EARLIER APPLICATION NUMBER: CIP OF 08/846, 247			
EARLIER FILING DATE: 1997-04-30			
EARLIER APPLICATION NUMBER: 60/119, 139			
EARLIER FILING DATE: 1999-02-08			
EARLIER APPLICATION NUMBER: 60/100, 880			
EARLIER FILING DATE: 1998-09-22			
EARLIER APPLICATION NUMBER: 60/087, 080			

EARLIER FILING DATE: 1998-05-28  
 NUMBER OF SEQ ID NOS: 34  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 20  
 LENGTH: 2401  
 TYPE: DNA  
 ORGANISM: Streptomyces venezuelae  
 US-09-320-878-20

Query Match 7.2% Score 179; DB 3; Length 2401;  
 Best Local Similarity 47.7%; Pred. No. 8.5e-30;  
 Matches 1018; Conservative 0; Mismatches 966; Indels 150; Gaps 11;

QY 206 CGCTCGGTTCCACATTCACCAACCTGCTCGAAGAGGAGCAAGTAGATGATGGGCAAG 265  
 DB 255 CCGTGGCAGAGACCTTGAGAGACACATGGCCAGACACTAGCGGAAGTATGGGCGCGG 314  
 QY 266 AGGCAATCGCTAAGAGTGGCATGTATCTCGGCCGCACTATCAACATCAAGCTCC 335  
 DB 315 ACGGTCGCGGCTCAACAGAGACATGGTCTGGGCCGAGATGATGACATCAATCCGGTGC 374  
 QY 326 CTCGCGTGGACGCTGCTGAGTGCATGTGTGAGATGCTGCTCGGCGGCTTGGGAG 385  
 DB 375 CGCACGGCGGCGGACCTAGAGACCTTCACAGAGAGCCCTGCTCTCTCGCGCACG 434  
 QY 386 CTGCGGCTCTATCCGCGGCAATTCAGAGCACTGAGTGCAGGCTACGATCAACACTTTT 445  
 DB 435 CGGTGCGGCAAGTCAAGAGGATCCAGGGTGGGGTGTGATGACACAGGCAACGACTTGG 494  
 QY 446 TGTGCAATGATCAGAGAGACAGCGCATGATGTGACAGACATGCTGCAGGAGCGGCTC 505  
 DB 495 CGGCAACCAACAGGAGAAACACCGCTTCCTCGTGAACGCCAATGTGAGACAGACGCG 554  
 QY 506 TCCGTGAATCTAGCACTCCGCTCCAGATGCTGTGAGACTCCACGCGGGTGGCT 565  
 DB 555 TCCGCGAGATGAGTCTCCGCGCTTCGAG---GCGTCTCCAAAGCCGCGGCGGCTCT 611  
 QY 566 TCATGACGGGTACATGAGCATGATGCGTGTGTCAGAGAGAACCTTAAATATCTTG 625  
 DB 612 TCATGTGTGCTCAACGCGCTCAACGGAAGCGTCTCGGGAACGAGACGACTCTCTCA 671  
 QY 626 ATGGAGTCTTCGAAAGATGGGGTGGGATGGCTATATCATGAGAGACTGTAGGCA 685  
 DB 672 ACAAGTGTGCGCACGACGATGGGGCTTCAGAGGCTGGGATGTGCTGACTGTGCTGCA 731  
 QY 686 CATACAGTACACAGAGAGCTGTGAGAGGCTTCGACCTCGAGATGCCGAGACTTCAC 745  
 DB 732 C---CCGGGACCGAGGACCATACCAAGGCGCTCGACACAGGAGATG-----775  
 QY 746 GCTTCGAGGAGAAACACTCAAGTTCAAGTCTTCCAAAGGAAAGCCCTTATCCAGTCA 805  
 DB 776 -----GGGTGTCAGACTCCCGCGGACGTCGCCGAAGGCGAGCCCTTCGCCCGGCA 827  
 QY 806 TTGACAGAGGGGTAGGAGTCTTCAAGTTCGTCAGAGAGTGTGCTCTCCGGAGTGA 865  
 DB 828 AGTTCTTCGGGA---GGCGCTGAAGAGCGCTCTCTGAAGGCGACGCTCCCGAGGCG 884  
 QY 866 CGGAGAACGCGCCGAGAGCACTGTCAACAACACCCCGAAGAGCGAGCTCTCTCGGA 925  
 DB 885 CCGTGAAGCGGTGGGAGGAGTCTCGGCCAGATGAGAGAGTTCGCTGCTCTCG 944  
 QY 926 AGTTGGCAAGAGGAGTGTGCTGTGAAGAGAGAAACAGTTTGCCTTGAAGA 985  
 DB 945 CCACTCCGGGCGCGCGGAGAGCGACAAAGGCGGCTGCCAGGGGTGTCCCGCAAG 1004  
 QY 986 AGAAGAGAGAGAGCTGATGTGCGCCCAAGCGCAAGCGACATACATACAGGGGAG 1045  
 DB 1005 TCGCGGAAGAGCGCGGTGCTCTGCGCAACAGAGGCGCGCTTCGCGCTGCGG 1064  
 QY 1046 GCTTCGCGCACTGAGGCGCTTACTACGACACTCCCTTGAAGGCGCTCAGCAAGCAG 1105  
 DB 1065 ACGCCGCAAGAGCATGCGGCTCATGCGCCCGACAGGCGCGCTGAC-CCCAAGGTACCGCG 1123

QY 1106 TCGAGAGCGCGCCATGCTACACCGTGGCGCTTACACACCGTCTCCCATTTAGGCG 1165  
 DB 1124 CTGGGAGCGCCACAGTCTGCTCCGAGCTCGGCGGCGCCCATCATACACATCAAGGCC 1183  
 QY 1166 AGAGTGTCTACGCGCGGAGCGGCTCCGGGATGGCTGAGAGGCTTCAACAGAGCCC 1225  
 DB 1184 CCGCGCGGTGGCGGTGCGAGGTGAGTACGAGAGCGGTTGAGAGACCTTCGGGAGCG 1243  
 QY 1226 CTGTACCCCTTAACCGCGAGCATATGACAGACTCTTCTTCAACAAAGACATGACCC 1285  
 DB 1244 ATCCCGCGCGGAGACTTCAGC-----CGCGCTTCACAC 1277  
 QY 1286 TGTGTGACTACTACCAACCCCAAGGCGGAGACAGCTGTACCGCATGAGGCGACT 1345  
 DB 1278 AGGCGCACAGCTCGAGCGGCGGCAAGCGGCGGTGTGACAGGAGGAGCTGACCGTGC 1337  
 QY 1346 ACACCGCGCAGAGAGACTGACCTACGAGACTGGCTGTGCTGTGCGGACAGGCAAG 1405  
 DB 1338 CCGCGCACGCGGAGTACCCATTCGCGGTCTGTCACCGGTGTACGCGCAGGTGAG- 1396  
 QY 1406 CGTACTAGACGACAGCTGCTGTGACAAAGCCACCAAGAGGTCCCGGCGATGCT 1465  
 DB 1397 -----CTCGGCAAGCACA 1409  
 QY 1466 TCTTGGCTCCGCCACCGCGAGAGAGCGCGCATCATCTGTCAGAGGCAACAGT 1525  
 DB 1410 CCATGAGGCGGTCAGTGTCTAGCGCAAGTGTAGAGCGCGCTGCTCAAGTGTACAGG 1469  
 QY 1526 ACAATGTTCAAGATCAGTGTGCTGCGCACCCACTACACCTTCAAGGCGAGACATCG 1585  
 DB 1470 GCACGACCAAGCT-----CAGCATCTCGGGCTTGGGATGA 1505  
 QY 1586 TCCCGCGGACGCGCTCCCTCGCGTGGCGGCTGCAAGGTCTATACAGACAGCGGAAA 1645  
 DB 1506 GTGCCACCGCGCTCTCTCGAGAGTGGCTGGGTACAGCGCGCGCGGCGGAGCA 1565  
 QY 1646 TCGAAAAGTCTGCGCTCGCCCAAGAGACAGCAGGTCAATCTGCGCGGCGCTTTA 1705  
 DB 1566 TCGCGAAGGCGGTGAGTGGCGGAGAGCGCGTACGCGGTCTGCTGCG-----1617  
 QY 1706 AGCCGACTGGGAGACCGAGGCGCGGAGCGCGGAGCATGAAAGCTCCCGCGGTGCG 1765  
 DB 1618 -CTACACAGAGGACCGAGGCGGTGACGCTTCGCAACCTGTGCGCGGTGACGAG 1676  
 QY 1766 ACCAGCTATTCCGAGCGGCGCGCGGAGACCAACCAACCTGCTGCTGATCAGCAGCG 1825  
 DB 1677 ACAAGTGTATCTGCGGTGTGCGGAGAGCCAGACCAACGATCGGTCTCAACACCG 1736  
 QY 1826 GCACCCCGAGAGATGCCCTGAGCTGACGCCAGCGCGCGCTCATTCAGCGCTGTAG 1885  
 DB 1737 GTTCGTGCTGTGATGAGCGGTGCTGTCCMAAGACCGCGCGGTCTGCGCATGTGTACC 1796  
 QY 1886 GCGGCAAGAGAGCGGCACTCATTTGCCAGCTGTGCTTGGGCACTTCAACCCCTGG 1945  
 DB 1797 CCGGCGAGCGGCGCGGAGGCGACCGCGCGCTGTCTACGCTGACGTCAACCGGAGCG 1856  
 QY 1946 GCAAGCTGTCTCCTCACTTCC-----AAGCGCTGAGAGCAACCCCGGCTTC 1996  
 DB 1857 GCAAGTCAACGAGAGCTTCCGCGCGCGGAGAACAGAGAGCGGCGCGGAGCCGA 1916  
 QY 1997 TCACTTCCGACCGAGCGCGGCGGAGCGCTGTAGCGGAGAGAGTCTACGTGAGTACA 2056  
 DB 1917 CAGAGTACCGCGGCGGTGACCAACAGCAACGATCGGAGGAGCATTCACGTGGGTACC 1976  
 QY 2057 GGTACTAGAGTGTGCGCAAGGAGCTCAATTTCCCTTGGGCAAGGCTGTCTCA 2116  
 DB 1977 GCTGTGACAAAGAGAGCTCAAGCGCTGTCCGTTGCGGACAGGCTGTGTGTA 2036  
 QY 2117 CCAGTTTCCCTTTTCAATCTCTCGGTCTCACAAAGAC---GGCAAGTACAGCTGT 2173  
 DB 2037 CCGTGTACGAGAGGCGCGGACCGTGTGTGATGCTGACAGGAGTGTGTGAAGGTCA 2096  
 QY 2174 CCCTCTCCGTGAAGAACACCGGCTCGGTGCGCGGACAGAGTGGCCCAAGCTCTACGTCA 2233

Accession	Sequence	Position
Db	CGGTCACGGTCCGCAACAGCGGGAACCGGCGCGCCAGAGAGTCTCTCAAGCGTACCTCG	2156
Qy	2234 AGCCCTTCACAGCGGGCCACAGATTAAACGGCCCCGTCACAGAGCTCAAGGCTTCGCAAGG	2293
Db	2157 GTGCGACACCCGAAAGTACGCGCTCCCGACGCGAAGAAAGTCTGTGGCTACACGAAGG	2216
Qy	2294 TCGAACTGCAGCCGGGGAAGACCAAGCGGTAC	2327
Db	2217 TCTCGCTCGCCGCGGGCGAGGCGCAACACGCTAC	2250

### RESULT 3

```

Sequence#23, Application US/09105537A
Patient No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438051
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ. ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 2430
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-23

```

Query Match	7.0%;	Score 175.2;	DB 4;	Length 2430;
Best Local Similarity	47.6%;	Pred. No. 5.6e-29;		
Matches 1016;	Conservative 0;	Mismatches 968;	Indels 150;	Gaps 11;

QY	206	CGTCGGTTCCACATATCAACCAAACTCGCTCGCAAGAGGCGAGGTAAATATATGGGCGAAAG	2655
Db	333	CCCTGGCCAGCACCTTGGACGACACACATGGCCGACACTACGGCAGAGGTATGGGCCGG	4122
QY	266	AGGCATCGCTAAGAGTGGCCGATGTGATTCCTCGGCCCGGACTATCAACATGCACAGCTCCC	3255
Db	413	ACGGTCGGCGCCCTCAACCAGGACATGGTCTCTGGGGCCCGGATGATGATCAACATCCGGGGTGC	4727
QY	326	CTCTCGGAGAACGTGGCGTTGAGTGCATPVGSTGAGATTCGGTTCCTCGCGGGCTTGGGAG	3855
Db	473	CGCACGGCGCGCGGAACACTACGAGACCTTCACCGAGAGACCCCTGGTCTCTCGCGCACCG	5322
QY	386	CTGCGGCTCTCATCCGCGGCAATTCCAGAGCACATGGAAGTGACAGGCTACAGATCAACACTTTT	445
Db	533	CGGTGCGCCAGATCATCAAGGGCATTCACAGGTTGGGGTCTGATGACACAGCCAGACACTTGC	5922
QY	446	TGTGCAATGATCAGAGGACACAGGGGCATGATGTTGGACAGACATCTGTACGACGCGGGCTC	5050
Db	593	CGGCCAAACACAGAGAGAACACCGCTTCTCCGTGAAGCCCAATGTGTACAGAGAGACGCG	6522
QY	506	TCCGTAATATCTACGACATCCCGTTCCAGATTGCTGTGCGAGATCTCCACCGCGGTGCGT	5555
Db	653	TCCCGGAGATGAGATTCCCGCGTTGCGAG---GCGTCTCCAAAGCGCGGCGGCTCTCT	709
QY	566	TCATATCAGCGGCTTACATATGGCATTCATATGGCGTGTGCTGACGAGAACCCCTAAATATCTTG	6255
Db	710	TCATATGTGCTTACACAGCGGCTTCAACGGGAAGCCGTCTCGGCGCAACAGAGAGCTCTTCA	7659
QY	626	ATGGGATCTCTCGAAAGGAATAGGGGTTGGGATGGCTTAATCATATGAGGACTGTGTAAGCA	6855
Db	770	ACAACGTGCTCGCGCACGAGTGGGGCTTCCAGGCGCTGGGTGATGTCCGATCTGGCTCGCA	8259
QY	666	CATATAGTACACAGAACCCGTTGTGGCAGGCGCTCGACACTCGAGATCCCGGACACTCCAC	7455
Db	830	C---CCCGGACCGGACCCATCAACAGAGGCTCTGACACAGGAGATG-----	873

Oy	746	GCTTCCGAGGAGAAACATCAGTTCACAGGTCCTCCACAGGAAAGCCCTTATCCAGCTCA	805
Db	874	-----GGCGTCGAGCTCCCGGCGAGCTCCCAAGGCGAGCCCTCGCCGCGAGCCCA	925
Oy	806	TTGACGAGAGGGCTAGAGGAAGTTCCTTCACTTGCTGTCAGAAAGTGTGTCGCTCCGAGATGA	865
Db	926	AGTTCTTCGGCGA---GGCGCTGAGACGGCGTCTTGAAGCGACGGTCCCCGAGGCGG	982
Oy	866	CGAGAAACGGCCCCGAGAGACTGTTCACACACCCCGGAAACGGCAGCTCTCTCCCGGA	925
Db	983	CCGTCAGCGCGGTGGCGGAGCGGATGTCGGCCAGATGAGAAAGTGTGGTGTCTCTCG	1042
Oy	926	AGTTGGCAACAGAGGCAATGTCCTCTGAAACAGAGAAACAGATTCGCTCTTGAGCA	985
Db	1043	CCACTCCGGGCGCCGCGCCCGAGACCGGAGAACAGGGCGGTGCCAAGCGGTGCCGGAAG	1102
Oy	986	AGAAAGAAAGACGCTGATGTTGTGGCCCCAACGCCAGAGGCCACATACCAGGGCGAG	1045
Db	1103	TCGCGGAGAAACGCGCGGTGTCCTTCGTCGACAGAGGGCCAGGCCCTGCGCTGCGCGTG	1162
Oy	1046	GCTCTGGCGCACTACAGGGCCCTTACTACGAGTCACTCCCTTGACGGCCCTCAGCAAGAGC	1105
Db	1163	ACGCCGGAAGAGCATTCGCGGTATGTGGCCCGACGCGCTGAC-CCCAAGTTCACGGC	1221
-Db	1106	TCGAGAGCGCCGCAATGATACCCGTGGCGCTTACACCAACCGTTCCTCCATTCTAGCG	1165
Oy	1222	CTGGGCGAGCGCCACGTCGTCGCGGACTGCGGCGGCGCGCACTGACACCATCMAAGCC	1281
Db	1166	AGCAGTGCCTACAGCCCGACGGGTGCTCCGGGATCGCGCTGGAGGGTCTTCAACAGAGCCC	1225
Oy	1282	CGCGCGGTGGGTGGAGGTGAGTACGTACAGACGGGTGAGAGACTTCGGAGCGCAG	1341
Db	1226	CTGTACCCCTTAAACCGCGACACATTTGACGAGCTCTTTCACCAAGACGACATGCACC	1285
Oy	1342	ATCCCGCGGGGAAACCTTCAAC-----CCGGCGTTTCAAC	1375
Db	1286	TGTGTGAGCTACTACACCCCAAGCGCGAGACAGTGGTACGCCGACATGTGAGGGGCACT	1345
Oy	1376	AGGGCACACAGCTGAGCCCGGGAGGGCGGGGCGCTGTACGACGCACTGACCGCTGC	1435
Db	1346	ACACCGCGCAGAGAGCTGCACCTACGAGCTCGGCTGTGTGTGCGGGCACGGCGAAAG	1405
Oy	1436	CCGCGCAGGCGAGTACCGCATGCGGTCCTGTCGACACGGGTGTATCCACCGGTGGAG-	1494
Db	1406	CGTACGTAGACAGACAGCTGTGTGTGACAAACGCCAACAGAGGTCCCGGGCGATGCT	1465
Oy	1495	-----CTGGCAGCCACA	1507
Db	1486	TCTTCGCGCTCGCCACCCGCGAGAGACGGCGCGCATCATCTCGTCAAGGGCAACAGT	1525
Oy	1508	CCATCGAGGCGCGTACAGTGTACGGGAAGTAGCAGCGCGGCTCTCAAGCTGACCAAG	1567
Db	1526	ACAGTTCAAGATGAGTTCGGCTCCGCAACCACTACACCTTCAAGGGCGACACATCG	1585
Oy	1588	GACGACCAAGCT-----CACGATTCGCGGCTTCGCGATGA	1603
Db	1586	TCGCCGGCAGAGGTCCTCCGCTCGCGGTGCGACGAAGGTCAATTACACACAGGCGCAAA	1645
Oy	1604	GTCGCCACCCGCTCTCCCTGGAGCTGGGTGGGTAGCGCGGCGGCGGCGAGCGACGA	1663
Db	1646	TCGAAAAGTCCGTCCCTCGCCACAGAGACAGCACAGTATCATCTTGCCCGGGCTTGA	1705
Oy	1664	TCGCGAAGGCGGTGAGTTCGGCCCGGAAGGCCCGTACGGCGGTGCTTTCG-----	1715
Db	1706	ACCGGCACTGGGAGACCGGAGGGCGCGACCGCGCGAGTAAAGTCTCCCGGCTGTCTGG	1765
Oy	1716	-CTACGAGACGGCACCGGAGGGGCTGACCGTTCGCAACTGTCTGCGGGGTACGAGG	1774
Db	1766	ACACGCTCATTCGCGAGTGGCGCGCGCGCAACCAACACCGTGTGTCTATGCAAGCG	1825
Oy	1775	ACAAGCTGATCTCGGTGTGCGGAGCGCAACCCAGACACGATCTGTGTCTCTCAACACCG	1834
Db	1826	GCACCCCCGAGAGATGCTTGGCTGTGACGCCACGCCGCCCTGATTCAGGCGTGTGACG	1885

```

      1835 GTTGTGTCGTCGATGATCCCTGGGCTGTCTCAAGACCCCGGGTCTCGAGATGTGGTACC 1894
QY 1886 GCGGCACGAGAGCGGGGACCTTCATTCGCGACGCGCTTTGGGAGCTTAAACCCCTCG 1945
Db 1895 CGGGCCAGAGGGGGCGCGAGAGCCGCGCGCTGTCTGAGGAGCTTAAACCCGAGCG 1954
QY 1946 GCAAGCTGTCTTCAGCTTCCCG-----AAGCGCTGCAGAGAACCCCGCGTTTC 1996
Db 1955 GCAAGCTGCAGAGAGCTTCCCGCGCGCGAGAACCGAGACGGGTGCGCGGGAGCCGA 2014
QY 1997 TCAACTTCGCGACCGAGAGCGGGGCGCAGCTGTACGAGAGAGCTTACGTTCGGGTACA 2056
Db 2015 CAACTTACCGGGGCGGTGAGAACCGAGACGATACCGGAGGAGATCCAGCTGCGGTACC 2074
QY 2057 GGTCTCTGAGATTGGCCAGAGAGAGCTCAATTTCCCTTTGGCGACGGCTGTCCGACA 2116
Db 2075 GCTGTCTGACAGAGAGAGAGCTCAAGCCGCTGTTCCTGTCGGGACAGCGGCTGCTACA 2134
QY 2117 CCACTTTGCTTTCCAACTCTCTCGCTGTCTCAAGAGAC---GGCMACTGAGCGTGT 2173
Db 2135 CCTGTCTCAGCAGAGAGCGCCCGACGCTGTGTGCGTACGTCAAGGGGTGTGTAAGGTCA 2194
QY 2174 CCTCTCTCGTGAAGAACCGGCTCCGTGCGCGCGACAGGTGGCCAGCTTACGTCA 2233
Db 2195 CGGTCTAGGTCGCGCAGACAGCGGGGAGAGCGCGCGCGAGAGGTCTTCAAGCGCTACCTCG 2254
QY 2234 AGCCCTTCCAAAGGCGGCAAGATTAAAGCCCGCTCAAGAGAGCTTCAAGAGG 2293
Db 2255 GTGCTACCGCCGAGAGCTACAGGCTCCGAGGCGGAAGAAAGCTGTGTGCTTACAGAGG 2314
QY 2294 TCGAAGTCGAGCCCGGCGAGAGAGAGCGGTGAC 2327
Db 2315 TCTGCTCGCCGCGGCGGAGCGGAAGACGGTGAAC 2348

```

```

RESULT 4
US-09-105-537-3
: Sequence 3, Application us/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/105,537A
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 13613
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-105-537-3

```

```

Query Match          7.0%; Score 175.2; DB 4; Length 13613;
Best Local Similarity 47.6%; Pred. No. 8.4e-29;
Matches 1016; Conservative 0; Mismatches 968; Indels 150; Gaps 11;

```

```

QY 206 CGCTCGGTTCCACATTCACCAACTCTGCTCGAAGAGCGAGTAAAGATGAGGCAAG 265
Db 4664 CCGTGGCAGACCTTGCAGAGACCACTTGGCCGACAGCTACGCGAAGCTATGGGCGCG 4723
QY 266 AGGCCATCGCTAAGAGTGCAGATGTATCTCGGCCGACTATCAACATCAACGCTCC 325
Db 4724 ACGGTCCGCGCTCAACAGAGACATGTCTCTGGGCCGATGATGAACAACATCGGGTGC 4783
QY 326 CTCTCGGTGAGCGGTGAGTGCATTTGATGAGATTCGTCCTGCGGGGCTTGGGAG 385
Db 4784 CGCAGCGGGCGGAGACTAGAGACTTTCAGCGAGAGCCCTGTGTCTCTCGCGCACCG 4843

```

```

QY 386 CTGCGGCTCTCATCCGCGGCAATTCAAGAGCACTGGAGTGCAGGCTACAGTCAACACTTTT 445
Db 4844 CGGTCCCGCAGATCAAGGAGGCAATCCAGGGTGGCGTCTGTATGACCAAGCGGACAGACTTCG 4903
QY 446 TGTGCAATATATAGAGAGGACAGCGCATATATGTGAGAGCATTCGTTCACGAGCGGGCTC 505
Db 4904 CGGCCAAACCAACAGAGGAACAACCGCTTCTCGTAAACCCCAATGTTCAGAGAGAGACGC 4963
QY 506 TCCGAATCTTACGCACTCCCGTTCCAGATGTGTGTGAGATGCCAGCGGGTGGCT 565
Db 4964 TCCGCGAGATCCAGATTCGCCGGCTTCGAG---GCTCTCTCAAGGCGCGGCGGCTCTCT 5020
QY 566 TCATGAGCGCGTACCAATGGCATATGCGGTGTCTGACAGCGAGAACCTTAATATCTTG 625
Db 5021 TCATGTGCTCAACAGCGCTCAACGCGGAGAACCGCTCTCGCGGACAGAGCTCTCA 5080
QY 626 ATGGATGCTTGAAGAAGAAATGGGTGGGATGGCTTAATCATATGAGCGCATCTACGGA 685
Db 5081 ACAAGCTGTGCGCACGAGCTGAGGCTTCCAGAGGCTGTGGGTATGTCCAGCTGTCCGA 5140
QY 686 CATACAGTACACAGAGAGCGGTGTGTGACAGCGCTGACCTCGAGATGCGGAGCTCCAC 745
Db 5141 C---CCCGGGCACGCGAGCGCATTCACAAAGGCGCTGACCAAGAGATG----- 5184
QY 746 GCTTCGAGAGAAACATCAAGTTCACTGCTCAACGAGAAAGCCCTTATTCACGTCA 805
Db 5185 -----GCGCTGAGACTCCCGGCGAGAGCTCCGAGGCGAGCGCTTCGCGCGGCGCA 5236
QY 806 TTGACAGAGGGGTAGGAGAAATTTCTCACTGTCTCAAGAAAGTGTGCTCTCGGAGTGA 865
Db 5237 AGTTCTTGCGCGA---GGGCTGTAAGACGCGGCTCTTGAAGGACAGGCTCCCGAGGCGG 5293
QY 866 CGGAGAACGCGCCCGAGAGCACTGTCAACACACCCCGCAAGCGAGCTCTCTCCGGA 925
Db 5294 CGGTGACGGGTGTGGCGGAGCGGATGTGTGCGCAATGAGAAAGTTGGTGTGTCTCTCG 5353
QY 926 AGGTTGGCAACGAGGAGCATGTGCTGTGTAAGAGAGAACAACTTCTGCTTGAAGCA 985
Db 5354 CCACCTCGCGCGCGCGCGCCGAGCGAGCAAGGCGGCTCCGAGGCGGTGTCTCCCAAGG 5413
QY 986 AGAAGAAAGAAAGCGCTGATTTGTGCGGCCCAAGGCCAAGGCGCAATACCGAGGCGGAG 1045
Db 5414 TCGCCGAGAAAGCGCGGCTGTCTGTGCGCAACGAGGCGGCGCTCCGCGCGGCG 5473
QY 1046 GCTCTCCGCACTCAGGCGCTACTACGAGTCACTCCCTTGAAGGCGCTCAAGAGACAG 1105
Db 5474 AGCGCGGCAAGAGCAATCGGGGTATGTGGCCGACGCGCTCGAC---CCGAAGGTCAACGCG 5532
QY 1106 TCGAGAGCGCGCCATGCTACACCGTGTGCGGCTTACACACACGCTTCTCCATTTAGGCG 1165
Db 5533 CTGGGAGCGCGCCAGCTGTCTCCGAGCTGCGGCGGCGCCACTGTGACATCAAGAGCG 5592
QY 1166 AGAGTGTCTCACGCGCCGAGCGCGCTCCGCGCATGTGGGTGAGAGGCTTCAACGAGCCCC 1225
Db 5593 CGGCGGGGTGTGGGTGTGCGAGCGGTGACGAGACGCGGTGAGAGACCTTCCGAGCGAG 5652
QY 1226 CTGTGATCCCTTAACGCGCACATATGACGAGCTTCTTTCACAAAGAGCATATGACAC 1285
Db 5653 ATCCCGCGCGGAGACTTCAGC-----CCGCGTTTCAMCC 5686
QY 1286 TGTGTGACTACTACACACCCCAAGGCGGAGACAGCTGTGTACCGCGCATATGAGGCGACGT 1345
Db 5687 AGGGCCACAGCTCGAGCCGCGGAGGAGGCGGCGGTGTATACAGAGGACGCTGACCGTGC 5746
QY 1346 ACACCGCGCAGAGAGCTGACCTTACGAGCTGGGCTGTCTGTCTGCGGCAAGCGCAAGG 1405
Db 5747 CGGCGCAGCGGAGTACCGCATTCGCGGTCTGTCCAGCGGTGTATACGCGAGGTGACG- 5805
QY 1406 CGTACTAGACAGACAGCTGTGTGTGACAAAGCCACCAAGAGAGTTCGCCCGGCGATGCT 1465
Db 5806 -----CTCGGCGACGCACA 5818
QY 1466 TCTTGGCTCTCCGCAACCCGCGAGAGAGCGGGCGCATCATCTGTCTCAAGGCGCAACAGCT 1525

```

Db 5819 CCATCGAGGCGGTCAGTCTACGAGGAGGAGGCGGCGTCCCAAGTACGACCAAG 5878  
QY 1526 ACAAGTTCAAGATCGAGTTGCGGTCCGCAACCACTACACCTCAAGGGCGACACATCG 1585  
Db 5879 GCACGACAAAGCT-----CAGATCTCGGGCTTCGGGATGA 5914  
QY 1586 TCCCGGCGACGAGCTCCCTCCGCGTGGCGGCTGCAGGTCAATTGACGACGAGCGGAAA 1645  
Db 5915 GTGCCACCCCGCTCTCCCTGGAGCTGGGTGGTACGCGCGGGCGGCGGACGCGAGCA 5974  
QY 1646 TCGAAAAGTCCGTCGCTCCGCAAGAGACGACGACGATCATCTGCGCGGCGCTTA 1705  
Db 5975 TCGCGAAGGCGCGTGGAGTGGCGCGGAGGCGCGTACGCGGCTCTTCGCG----- 6026  
QY 1706 AGCGGACGAGGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1765  
Db 6027 -CTACGACGACGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6085  
QY 1766 ACCAGCTCATTTGCGGAGTGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1825  
Db 6086 ACAAGCTGATCTCGGCTGTGGGAGCGCAACCGGACGATGCTGTCTCAACACCG 6145  
QY 1826 GCACCCCGGAGAGATGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1885  
Db 6146 GTTCGTCGCTGATGAGCGGCTGTCTCAAGACCGCGGCTGTCTGAGCATGTGTACC 6205  
QY 1886 GCGGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1945  
Db 6206 CGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6265  
QY 1946 GCAAGCTGTCCTCAGCTTCCCG-----AAGCGCTGAGAGGACGCGGCGGCTTC 1996  
Db 6266 GCAAGCTGTCAGGAGCTTCCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6325  
QY 1997 TCACTTCGCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2056  
Db 6326 CAAGCTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6385  
QY 2057 GGTACTACGAGTTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2116  
Db 6386 GCTGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6445  
QY 2117 CCAGTTTTCCTTTCCTCAATCTCTCGGTCTCAAGAGGAGGAGGAGGAGGAGGAGG 2173  
Db 6446 CCGTCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6505  
QY 2174 CCTCTCCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2233  
Db 6506 CGGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6565  
QY 2234 AGCGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2293  
Db 6566 GTGCCAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6625  
QY 2294 TCGAAGTGAAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2327  
Db 6626 TCTCGCTCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6659

RESULT 5  
US-09-428-517-1/c  
; Sequence 1, Application US/09428517  
; Patent No. 6251636  
; GENERAL INFORMATION:  
; APPLICANT: Bellich, Mary C.  
; APPLICANT: Shah, Sanjay Krishnakant  
; APPLICANT: McDaniel, Robert  
; APPLICANT: Tang, Li  
; TITLE OF INVENTION: RECOMBINANT OLEANDROLIDE POLYKETIDE SYNTHASE  
; FILE REFERENCE: 30062-20029 00  
; CURRENT APPLICATION NUMBER: US/09/428,517  
; CURRENT FILING DATE: 1999-10-28

EARLIER APPLICATION NUMBER: 60/120,254  
EARLIER FILING DATE: 1999-02-16  
EARLIER APPLICATION NUMBER: 60/106,100  
EARLIER FILING DATE: 1998-10-29  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO: 1  
LENGTH: 50937  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA  
US-09-428-517-1

Query Match 5.5%; Score 138.6; DB 4; Length 50937;  
Best Local Similarity 54.0%; Pred. No. 8.6e-21;  
Matches 354; Conservative 0; Mismatches 289; Indels 12; Gaps 3;

QY 1710 CGACTGGAGACCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1769  
Db 3417 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3358  
QY 1770 GCTATTGCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1829  
Db 3357 CCGTATGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3298  
QY 1830 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1889  
Db 3297 CTCCCTACGATGCGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3238  
QY 1890 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1949  
Db 3237 CCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3178  
QY 1950 GCTGCTTCAGCTTCCCG-----CAAGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 2003  
Db 3177 GCTGAGCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3118  
QY 2004 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2060  
Db 3117 CTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3058  
QY 2061 CTAGGAGTTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2120  
Db 3057 GTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2098  
QY 2121 TTTTCCCTTTCCTCAATCTCTCGGTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGG 2177  
Db 2097 CTTCGACTACGCGGAGCTGAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2038  
QY 2178 CTGCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2237  
Db 2037 CACGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2078  
QY 2238 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2297  
Db 2077 GTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2018  
QY 2298 ACTGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2352  
Db 2017 GCTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2058

RESULT 6  
US-09-134-078-5  
; Sequence 5, Application US/09134078  
; Patent No. 6368844  
; GENERAL INFORMATION:  
; APPLICANT: Bylina, Edward J.  
; APPLICANT: McDaniel, Robert  
; TITLE OF INVENTION: GLYCOSIDASE ENZYMES  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Gray Cary Ware & Freidenrich LLP





Db 218227 GAATCGCGTGTGGC-----GTTGCTGGATTGAGCGGGGTGACAGCTCGAACCC 218274  
 QY 2308 GCGGAGACGAGCGCGTGCATATCGAGAGACAGAGACTGCTGCTGATTTTGTAT 2367  
 Db 218275 GCGCAGACTCGCGGGTAAAGATCGAGGGGCGCGGACTGCTGCGCCGCTACGAGGCG 218334  
 QY 2368 GAGG 2371  
 Db 218335 GAGG 218338

## RESULT 8

US-09-103-840A-1  
 ; Sequence 1, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; TITLE OF INVENTION: TUBERCULOSIS  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 4411529  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; OTHER INFORMATION: H37Rv  
 US-09-103-840A-1

Query Match 4.58; Score 111.6; DB 4; Length 4411529;  
 Best Local Similarity 52.08; Pred. No. 1.6e-14;  
 Matches 345; Conservative 0; Mismatches 289; Indels 30; Gaps 3;

QY 1723 GAGGCGCGGACCGCGAGACATGAGCTCCCGCGTGTGGACGACGATTCGCGGAC 1782  
 Db 217581 GAGGCGCTTCGACAGCGCGCATCTGTGCTGCATGGGTGATGAGATGCGCTGATGCGCGCA 217640  
 QY 1783 GTGGCGCGCGGACCCAAACCGCTGCTGCTATGACGAGCGGCAACCCCGGAGAGATG 1842  
 Db 217641 GTGCGGTGCGGACGAGATACCGTGTGTGCTTGAGAGCGGCAACCGCGGTGACCATG 217700  
 QY 1843 CCTGCGCTCAGCGGACGCGCGCGCTATCCAGGCGCTGATGCGGCGCAAGAGAGCGGC 1902  
 Db 217701 CCTGCGCGCTCAGCGGACGCGCGCTATCCAGGCGCTGATGCGGCGCGGCGGTGCGC 217760  
 QY 1903 AACTCATTGCGCGAGCTGCTTTGGGACTACACCCCTGCGGCAAGCTGCTCCTACG 1962  
 Db 217761 CAGCGCGCTTGGGAGATGTGACCGGCAAGTATCCTTCGGGCGCGGCGCGCATACAC 217820  
 QY 1963 TTCCCGCAGCGCGCTGAGCAACCCCGCTTCTCACTTCGCAACCGAGCGCGGCGC- 2021  
 Db 217821 TTCCCGGCTGATCTCGGTGAGACGACGCTCGCAACCGCGCGAGGTGCGCGCGGTG 217880  
 QY 2022 -----CAGCGTGTACGGGAGAGAGCTGCTACTGCGGTGACAGGTACTACGAG 2067  
 Db 217881 GGGACATGACACGATCGCTACACCGAGGCGCGGATGTTGTTACCGCTGTTGGC 217940  
 QY 2068 TTGGCGCAGAGAGCGATTCCTCCCTTTGGCGAGGCGCTGCTACACCACTTTTGGC 2127  
 Db 217941 AGCACAATACAGACCCCGATGTTGCGCTTGGTACAGGCTTGTCTTATACCACTTTGCGAG 218000  
 QY 2128 TTTTCCATCTCTCCGTGTCTACAAAGAGAGCGCAAGCTGATGCTGCTCTCTCCGTGAAG 2187  
 Db 218001 TATCGTACCTGTGTTGAGAGGCG---GGCGACCGGTGACGCGCATTTTCAACGTTACG 218057  
 QY 2188 AACACGCGCTCGTCCGCGGCGCAGAGTGGCCAGCTTACGTCAAAGCCCTCCAGGCG 2247

Db 218058 AACACGCGCGACCGCAGCGGCGGAGATGTCCCGACGTGTATATGATCCAGCTCCGCGC 218117  
 QY 2248 GCCAGATTAAACGCGCCCGCTCAAGGAGCTCAAGGCGCTTCGCAAGAGTCACTGACGCGC 2307  
 Db 218118 GAATGCGCGTTCG-----GTTGCTGGATTTCGAGCGGGTGTGACCTGCAACCC 218165  
 QY 2308 GCGGAGACGAGCGCGTGCATATCGAGAGAGAGAGAAATGATGCTGCTGCTATTTTGTAT 2367  
 Db 218166 GCGCAGACTCGCGGGTAAAGATCGAGGCGGACCGCGAGCTGCTGCGCCGCTACAGCGGC 218225  
 QY 2368 GAGG 2371  
 Db 218226 GAGG 218229

## RESULT 9

US-09-147-236-1  
 ; Sequence 1, Application US/09147236A  
 ; Patent No. 6316251  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TONOUCHI, Naoto  
 ; APPLICANT: TSUCHIDA, Takayasu  
 ; APPLICANT: YOSHINAGA, Fumihito  
 ; APPLICANT: TAHARA, Naoki  
 ; APPLICANT: HAYASHI, Takahisa  
 ; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE  
 ; FILE REFERENCE: 6537-011-0PCT  
 ; CURRENT APPLICATION NUMBER: US/09/147,236A  
 ; CURRENT FILING DATE: 1999-04-08  
 ; EARLIER APPLICATION NUMBER: PCT/JP97/03633  
 ; EARLIER FILING DATE: 1997-10-09  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 16836  
 ; TYPE: DNA  
 ; ORGANISM: Acetobacter xylinum

FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: (869)..(1891)  
 FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: (3101)..(5368)  
 FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: (5373)..(7778)  
 FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: (7784)..(11761)  
 FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: (11764)..(12231)  
 FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: (12448)..(14652)  
 FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: (12448)..(14652)  
 OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or  
 OTHER INFORMATION: c  
 US-09-147-236-1

Query Match 4.18; Score 101.4; DB 4; Length 16836;  
 Best Local Similarity 50.58; Pred. No. 6.6e-13;  
 Matches 246; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 242 AGCGAGTAAGTATGATGGGCAAGAGCGCATCGCTAAGAGTGGCGCATGTGATCTCGGCGC 301  
 Db 12845 AGCGCGGTGCTATGATCGGCGCGGCGGAGCATGCGAGAGCGGCTTCAACATCTGCTGGCG 12904  
 QY 302 CGACTATCAACATCGAAGCGTCCCTCTCGGTGAGCGTGGCTTCGAGTGTGAGGAG 361  
 Db 12905 GCGGTGCGGACCTGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 12964  
 QY 362 ATCGTTCCTGCGGCGGCTTGGGAGCTGCGGCTTCATCGCGGCGCATTCAGAGCACTGGAG 421



Db 12965 ATCCGCTGCAGACCGGGCGCATGTTGGGACAGCACCATTGAGGGCTGACATCGCAGCATG 13024  
QY 422 TGCAGGCTACGATCAAGCATTCTTTTGTGCAATGATCAGAGGAGCAGCGCATGATGTCG 481  
Db 13025 TGATCTCCAGCCTACAGCATTTATGCGATGATGACCTCGAAGCCTCGCATGACCATG 13084  
QY 482 AGAGCATGTCACGAGCGGGCTCTCCGTGAATCTACGACCTCCGCTTCAGATTGCTG 541  
Db 13085 GCGCGATATGACCTCTGTGTCATGCGTGAAGACGACCTGCTGGCTGAGATCGCGC 13144  
QY 542 TGCAGACTCCAGCCGGGTCTTCATGACGAGGTACAAATGATGAATGAGGCTGCTG 601  
Db 13145 TTGAACCGGGGATCCGGGGCGGTGCTGCTGTAACACCGGCTACAGCATTATG 13204  
QY 602 GCAGCGAAGACCTTAATATCTTGTGATGATGCTTCGAAAGAAATGGGTTGGGATGCCC 661  
Db 13205 CGTGTGAAGAACCCGTACTGCTGTAACAAGAGCTGAAGAGGACTGGCATTTATCCGGCT 13264  
QY 662 TAATCATGAGCGACTGTGACGACATACAGTACCACAGAGCCGTTGGCAGGCTCG 721  
Db 13265 TTGTCATGTCGACTGGGGGGGCGACAGCATTCCTCCGCGGGCGCGCTGGCGGCTGG 13324  
QY 722 ACCTCGA 728  
Db 13325 ATCAGGA 13331

## RESULT 10

US-09-147-236-10  
; Sequence 10, Application US/09147236A  
; Patent No. 6316251  
; GENERAL INFORMATION:  
; APPLICANT: TONOUCHI, Naoto  
; APPLICANT: TSUCHIDA, Takayasu  
; APPLICANT: YOSHINAGA, Fumihito  
; APPLICANT: TAHARA, Naoki  
; APPLICANT: HAYASHI, Takahisa  
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE  
; FILE REFERENCE: 6537-011-0PCT  
; CURRENT APPLICATION NUMBER: US/09/147,236A  
; CURRENT FILING DATE: 1999-04-08  
; EARLIER APPLICATION NUMBER: PCT/JP97/03633  
; EARLIER FILING DATE: 1997-10-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 16836  
; TYPE: DNA  
; ORGANISM: Acetobacter xylinum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1891)..(2922)  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence is the same as SEQ ID NO:1  
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or  
; OTHER INFORMATION: t  
; OTHER INFORMATION: t  
US-09-147-236-10

Query Match 4.1%; Score 101.4; DB 4; Length 16836;

Best Local Similarity 50.5%; Pred. No. 6.6e-13;

Matches 246; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 242 AGCAGTAAGATGATGGGCAAGAGCCATCGCTAAGAGTGGCATGTGATCTCGGCC 301  
Db 12845 AGCGCGGTCTCATGATCGGGCGAGGCGATGCGAGGCGCTTCAACATCTCTGGCG 12904  
QY 302 CGACTATCAACATGCAAGCTCCCTCTCGGAGAGTGGGAGTGGGATGAGG 361  
Db 12905 GCGGTCCGAGCTGACGCGGACCGCGTGGGCGCCGCACTTTGAATATGCGGCGAAG 12964  
QY 362 ATCCGTTCTGGGCGGCTTGGGAGCTGCGCTCTCATTCGCGGCAATTCAGAGCACTGGAG 421

Db 12965 ATCCGCTGCAGACCGGGCGCATGTTGGGACAGCACCATTGAGGGCTGACATCGCAGCATG 13024  
QY 422 TGCAGGCTACGATCAAGCATTCTTTTGTGCAATGATAGAGAGCAGCGCATGATGTCG 481  
Db 13025 TGATCTCCAGCCTACAGCATTTATGCGATGATGACCTCGAAGCCTCGCATGACCATG 13084  
QY 482 AGAGCATGTCACGAGCGGGCTCTCCGTGAATCTACGACCTCCGCTTCAGATTGCTG 541  
Db 13085 GCGCGATATGACCTCTGTGTCATGCGTGAAGACGACCTGCTGGCTGAGATCGCGC 13144  
QY 542 TGCAGACTCCAGCCGGGTCTTCATGACGAGGTACAAATGATGAATGAGGCTGCTG 601  
Db 13145 TTGAACCGGGGATCCGGGGCGGTGCTGCTGTAACACCGGCTACAGCATTATG 13204  
QY 602 GCAGCGAAGACCTTAATATCTTGTGATGATGCTTTCGAAAGAAATGGGTTGGGATGCCC 661  
Db 13205 CGTGTGAAGAACCCGTACTGCTGTAACAAGAGCTGAAGAGGACTGGCATTTATCCGGCT 13264  
QY 662 TAATCATGAGCGACTGTGACGACATACAGTACCACAGAGCCGTTGGCAGGCTCG 721  
Db 13265 TTGTCATGTCGACTGGGGGGGCGACAGCATTCCTCCGCGGGCGCGCTGGCGGCTGG 13324  
QY 722 ACCTCGA 728  
Db 13325 ATCAGGA 13331

## RESULT 11

US-08-387-942C-1  
; Sequence 1, Application US/08387942C  
; Patent No. 5939289

; GENERAL INFORMATION:

; APPLICANT: ERTESVAG, HELGA

; APPLICANT: VALLA, SYBEN

; APPLICANT: SKJAK-BRAEK, GUDMUND

; TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES

; TITLE OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP

; STREET: P.O. BOX 747

; CITY: FALLS CHURCH

; STATE: VA

; COUNTRY: USA

; ZIP: 22042

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/387,942C

; FILING DATE: 09-May-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: MURPHY JR, GERALD M.

; REGISTRATION NUMBER: 28,977

; TELECOMMUNICATION INFORMATION:

; REFERENCE/DOCKET NUMBER: 1809-106P

; TELEPHONE: 703-205-8050

; TELEFAX: 703-205-8050

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12588 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Azotobacter vinelandii

; STRAIN: E

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 290..1951  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2227..6438  
 NAME/KEY: CDS  
 LOCATION: 6702..9695  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 9973..12588  
 US-08-387-942C-1

Query Match 3.4%; Score 84.6; DB 2; Length 12588;  
 Best Local Similarity 43.3%; Pred. No. 2.5e-09;  
 Matches 451; Conservative 0; Mismatches 584; Indels 6; Gaps 1;

QY 1197 CATGCGTGGAGGCTCTTCAACGAGCCCTGGTACCCCTTAACCCGACACATTGACGA 1256  
 DB 2601 CATCCCCGCGACGAGCGCGCCGATGCGACGCTGAGCCGGTGGAAATCCGCGA 2660  
 QY 1257 GCTCTTCTTACCAAGACGACATGACCTGTGTGACTACTACACCCCAAGGCGGACGA 1316  
 DB 2661 GATGTCCGGGTAGGTTTGACACCCACAGACACATCAACCTGACATCCGCGACAG 2720  
 QY 1317 CACGTGTAGCGCCGACATGAGGCGACGTAACCCGCGACGAGACTGCACTAGAGCT 1376  
 DB 2721 CGTGGCCCGACGACAAACGCTGACGCGCTTCTGCGCCGACTACAGGTGGCGGGTGT 2780  
 QY 1377 CGGCTCTGTCTGTGGGACCGCAAGGCGTACGTAACGACGACGCTGTGCTGACAA 1436  
 DB 2781 CGAGAAACAGCTCTGTAACAAACAGACCGCGACGCTTCAACATGTCACAGACCA 2840  
 QY 1437 CGGCAACAGAGGTCCTCGCGGATGCTTCTTGGACTCGGACCCGCGAGAGACGG 1496  
 DB 2841 CGACTTCTGTCTGAGCAACAGTGCCTTACGCGCAAGCGCGCGCTGTGTGTGA 2900  
 QY 1497 CGGATCAATCTGTCTAGGCGCAACAGCTAAGATTCAAGATCGATCGCTCGCAC 1556  
 DB 2901 GCGCGGCTGTGATGACCTCCCTCAATCCATGACATCTGATGACGCGGCGCTACTA 2960  
 QY 1557 CACCTACACCTCAAGAGGAGACACATCTCCCGCGGACGCTCTCCGCTGGGG 1616  
 DB 2961 CGACAAAGGCTTGAAGGCTGACAGTGAATGAGCCCAAGACGACGCTTCCAGAAC 3020  
 QY 1617 CTGCAAGGCTATGACGACGAGCGAATCGAAAGTCCGCGCTCGCCAGAGCA 1676  
 DB 3021 CGAGATCTAGGCGAAGCGGCTGTACGGGGTGGCGTCTACGCGCGCCAGAGCTCA 3080  
 QY 1677 CGACGAGTATCTGTGCGCGGCTTACGCGGATGAGGAGCGAGGCGCGGACCG 1736  
 DB 3081 CTTGCAACAGATTCAGACAAATTCGAGAAAGCGGCTTATGCGAAGTCTCTGCA 3140  
 QY 1737 CGCAGCATAGAGTCCCGGCGTCTGTGACAGCTCAATTCGCGAGCGGCGCGGAA 1796  
 DB 3141 GTCTACAGACGACCGCGGGGTGTCCGGCACTTTTACGTACACCGGACCTTGCT 3200  
 QY 1797 CCGAAACACGCTGTGTGATGAGAGGCGACCCCGAGAGATGCTGTGTGAGCG 1856  
 DB 3201 CGAAGGCAAGCTATCAGCGGCTCGGCAATTCACATACGCGCATTCAGAGCGGCG 3260  
 QY 1857 CAGCGCGCGGCTATCCAGGCTGTAGGGGCGCAACAGAGCGGCACTGATCCGA 1916  
 DB 3261 CGGACCGGCTACAGACGCTTACGCAACAGCTGACGCTGTGCGAGCGGGCGGT 3320  
 QY 1917 GCTGCTTGTGGAGTACCAACCCCTCGGCAAGCTGTCTCAAGCTTCCCAAGCGCT 1976  
 DB 3321 ACGGCTGTATGGCGCACTGACGCTTCCAGCAGTCCGCGAGTGGCCAGCGAGC 3380  
 QY 1977 GCAGAGCAACCCCGGCTTCTCA-----CTTCGACCGAGCGCGGCGCAGCTGTA 2030  
 DB 3381 CCTCAAGGACGCGGCGGCGACAGATGCGCTGAGGGGAGCGAGGCGCCAGAGAGCTGCT 3440

QY 2031 CGGCGAGACGCTACGTCGGGTACAGTACTACAGATTGCGCAAGAGCTCAATT 2090  
 DB 3441 CGGCGAGCGCGGCGACGACCGGCTTACAGCGGATCCGCGCAACGACATCTCGAGCGG 3500  
 QY 2091 CCGCTTGGCCACGCGCTGTCTTACACGATTTTCCCTTTTCAATCTCTCGTCTCA 2150  
 DB 3501 GGCAGGCGCGCAACACTGACCGCGCGCGCGCGGCGACACTTCCGCTTCTCCGCGG 3560  
 QY 2151 CAAGGACGCGCAAGCTGACGCTGTCTCTCTGTAAGAAACACCGGCTCCGCGCGCG 2210  
 DB 3561 CACGACAGCTACGCGCACGAGCGCGCACTTCAACGACCTGATCAACCGACTTTCAGCG 3620  
 QY 2211 ACAGGTGCGCGGCTCTAGCT 2231  
 DB 3621 CGAGAGGACAGACATCGACT 3641

RESULT 12  
 US-09-197-649-7  
 ; Sequence 7, Application US/09197649  
 ; Patent No. 6194550  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gold, Larry  
 ; APPLICANT: Tuerk, Craig  
 ; APPLICANT: Pribnow, David  
 ; APPLICANT: Smith, Jonathan D.  
 ; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation  
 ; FILE REFERENCE: NEX02/CI-CON  
 ; CURRENT APPLICATION NUMBER: US/09/197, 649  
 ; EARLIER FILING DATE: 1998-11-23  
 ; EARLIER APPLICATION NUMBER: 07/829,461  
 ; EARLIER FILING DATE: 1992-01-31  
 ; EARLIER APPLICATION NUMBER: 07/739,055  
 ; EARLIER FILING DATE: 1991-08-01  
 ; EARLIER APPLICATION NUMBER: 07/561,968  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 7  
 ; LENGTH: 390  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Sequence  
 ; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed  
 ; OTHER INFORMATION: fragments having NcoI restriction sites.  
 US-09-197-649-7

Query Match 3.4%; Score 84.4; DB 4; Length 390;  
 Best Local Similarity 51.9%; Pred. No. 1.2e-09;  
 Matches 190; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 1323 GTAGCGGACATGAGGAGGACGTACACCGCGCGAGAGATGACCTACAGCTCGGCT 1382  
 DB 8 GGAGGACGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 67  
 QY 1383 GGTGCTGTGGGCGACGCAAGGCGTACGTAGACGACGACGACGCTGTGTGCAACGCGAC 1442  
 DB 68 CGAGGACGAGACGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 127  
 QY 1443 CAGCAGGTCCCGGCGATGCTTCTTGGCTTCCGCAACCCGAGAGAGAGCGGCGCAT 1502  
 DB 128 CGAGGACGAGACGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 187  
 QY 1503 CAATCTGTCAAGGCGCAACGTAAGTTCAAGATCGAGATTGCGCACCCACCTA 1562  
 DB 188 CGAGGACGAGACGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 247  
 QY 1563 CACCTTCAAGGCGACACATGCTTCCCGCGCACGCGCTCTTCCGCTGAGCTGCA 1622  
 DB 248 CGAGGACGAGACGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 307



```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/625,140
FILING DATE: 10-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Dillahunty, T. Gene
REGISTRATION NUMBER: 25,423
REFERENCE/DOCKET NUMBER: 010055-056
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-7400
TELEFAX: 415-854-8275
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
FEATURE:
NAME/KEY: CDS
LOCATION: join(311..375, 446..2205, 2270..2679)
FEATURE:
NAME/KEY: Intron
LOCATION: 376..445
FEATURE:
NAME/KEY: Intron
LOCATION: 2206..2269
US-08-462-090-1

Query Match          3.3%; Score 83.8; DB 3; Length 3033;
Best Local Similarity 50.1%; Pred. No. 2.7e-09;
Matches 264; Conservative 0; Mismatches 257; Indels 6; Gaps 2;

QY 214 TCACATTCAACCAACTGCTGCTGAGAGGAGGAGTGAATGATGGGCAAGAGCCATC 273
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 720 TCACAGTGGAGATGTCATTTGATCCCGAAGGAGGACAGTTCATGAGAGGAG 779
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 274 GCTAAGAGTGGAGATGATGATCCCGGCC--GACTATCAACATGCAAGCTCCCTTC 330
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 780 GCTCGGGGATTCATGTCATCTGCTGCTGCTGGCCCTGGGAAGAGCTCCGAG 839
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 331 GGTGAGAGTGGCTTCGATGATGTTGGTGAAGATCCGTTCTCGGGGCTTTGGAGCTGC 390
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 840 GCGGATCGGAAGCTGGAGGGCTTCGGTGTGATTCATCTCAGCGGCAATTCGATGGGT 899
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 391 GCTCTATCCGGGCAATTCAGAGCACTGAGTGAAGCTAATCAACGACTTTTGTGC 450
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 900 CAATCCATCAACGCGATCCAGTGGAGGCTGAGGCGACGCAAGCACTATATCCTC 959
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 451 AATGATCAGGAGGAGGAGGAGATGATGTCAGAGCATGTCAGGAGGCGGCTCCGCT 510
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 960 AATGATCAGGAGGAGGAGGAGATGATGTCAGAGCATGTCAGGAGGCGGCTCCGCT 510
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 511 GAATCTAGCACTCCGCTTCAGATTGCTGTGCGAGACTCCAGCGGGTGCCTTCATG 570
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1020 GACTCTATACCTTGGCCATTTGGCCAGCGGGTT--AGGCCAATGTGCTTCTGTGATG 1076
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 571 ACGGCTCAAGTGGATGATGGCGGTGCTGTCAGCGAGCAACCCCTAATATCTTATGGG 630
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1077 TGCTCTGAACAAAGGTCAATACCACTTGGGCTGCGAGATCACTACAGCTGAGACT 1136
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 631 ATGCTTGAAGGAATGAGGGTGGGATGCGCTAATCATGAGAGGAGTGTGATGCGCAATAC 690
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1137 GTGCTAAAGACACGCTGGGGTCCAGGCTATGTCATAGCAGCTGGGAACGACACAGCAC 1196
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 691 AGTACCAAGAACGCTTGTGAGGAGCTGAGACTGAGATGCGCCG 737
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1197 AGCAGTGTCAAAAGCGGCAATTCGTGGGCTTGAACATGTCAATGCTGCG 1243
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

RESULT 15
US-08-463-461-1
Sequence 1, Application US/08463461
Patent No. 6103464
GENERAL INFORMATION:
APPLICANT: Fowler, Timothy
APPLICANT: Barnett, Christopher C.
APPLICANT: Shoemaker, Sharon
TITLE OF INVENTION: Saccharification of Cellulose by Cloning
TITLE OF INVENTION: and Amplification of the Beta-glucosidase Gene of
TITLE OF INVENTION: Trichoderma Reesei
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A.
ZIP: 94304-1013
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,461
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Christopher L. Stone
REGISTRATION NUMBER: 35,696
REFERENCE/DOCKET NUMBER: GC78D4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-846-7555
TELEFAX: 415-845-6504
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3033 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
FEATURE:
NAME/KEY: CDS
LOCATION: join(311..375, 446..2205, 2270..2679)
FEATURE:
NAME/KEY: Intron
LOCATION: 376..445
FEATURE:
NAME/KEY: Intron
LOCATION: 2206..2269
US-08-463-461-1

Query Match          3.3%; Score 83.8; DB 3; Length 3033;
Best Local Similarity 50.1%; Pred. No. 2.7e-09;
Matches 264; Conservative 0; Mismatches 257; Indels 6; Gaps 2;

QY 214 TCACATTCAACCAACTGCTGCTGAGAGGAGGAGTGAATGATGGGCAAGAGCCATC 273
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 720 TCAGAGTGGATGTCATTTGATCCCGAAGGAGGAGTTCATGAGAGGAGTGAAG 779
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 274 GCTAAGAGTGGAGATGATGATCCCGGCC--GACTATCAACATGCAAGCTCCCTTC 330
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 780 GCTCGGGGATTCATGTCATCTGCTGCTGCTGGCCCTGGGAAGAGCTCCGAG 839
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 331 GGTGAGAGTGGCTTCGATGATGTTGGTGAAGATCCGTTCTGGCGGGCTTGGAGCTGC 390
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 840 GCGGATCGGAAGCTGGAGGGCTTCGGTGTGATTCATCTCAGCGGCAATTCGATGGGT 899
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```



